

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2005, 17:16:07 ; Search time 162 Seconds
(without alignments)
128.920 Million cell updates/sec

Title: US-10-644-927-1

Perfect score: 304

Sequence: 1 KTYGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTEH 54

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Genesecp1980s:*
- 2: Genesecp1990s:*
- 3: Genesecp2000s:*
- 4: Genesecp2001s:*
- 5: Genesecp2002s:*
- 6: Genesecp2003as:*
- 7: Genesecp2003bs:*
- 8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69.5	22.9	420	6	ADA33974 Acinetoba
2	64	21.1	61	2	AAW17990 Sakacin P
3	62.5	20.6	558	8	ADR89438 crybun2-o
4	62	20.4	42	2	AAW11633 Lactobaci
5	60	19.7	37	2	AAW66446 Catonic
6	60	19.7	37	3	AAW91745 Cationic
7	60	19.7	300	2	AAW56481 CD38. 3/2
8	60	19.7	300	2	AAW37928 Amino aci
9	60	19.7	300	4	AAW89069 Human CD3
10	60	19.7	300	5	ABG61817 Prostata
11	60	19.7	300	6	ABU09711 CD38 prot
12	60	19.7	300	6	ABU09714 CD38 prot
13	60	19.7	300	7	ADD18904 Human dis
14	60	19.7	300	7	ADD45262 Human pro
15	60	19.7	300	7	ADN40030 Cancer/an
16	60	19.7	300	8	ADM32177 Human CD3
17	60	19.7	300	8	ADQ59526 Human can
18	60	19.7	300	8	ADQ88192 Human 903
19	60	19.7	688	4	ABG06883 Novel hum
20	59.5	19.6	513	2	AAW79755 Euphorbia
21	59	19.4	44	2	AAR14564 Bacterioc
22	59	19.4	44	2	AAR91745 Brevicin,
23	59	19.4	62	2	AAR25078 PA-1 bact
24	59	19.4	300	6	ABU09712 CD38 prot
25	59	19.4	399	8	ADN23488 Bacterial

ALIGNMENTS

RESULT 1

ADA33974
ID ADA33974 standard; protein; 420 AA.

XX AC ADA33974;

DT 20-NOV-2003 (first entry)

XX AC Acinetobacter baumannii protein #1135.

XX KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
XX KW plant biocontrol agent.

XX OS Acinetobacter baumannii.

XX FN US6562958-B1.

XX PD 13-MAY-2003.

XX PF 04-JUN-1999; 99US-00328352.

XX PR 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton G, Bush D;

XX DR WPI; 2003-576092/54.

XX N-PSDB; ADA29848.

XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX PT for diagnosing a bacterial disease, as components of antibacterial
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX plants.

XX Example; SEQ ID NO 5261; 328pp; English.

XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents
XX CC for diagnosing a bacterial disease, as components of antibacterial
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of
XX CC A. baumannii and other Acinetobacter species in a sample, in screening
XX CC compounds for the ability to interfere with the A. baumannii life cycle
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for an A.
XX CC baumannii protein.

XX SQ Sequence 420 AA;

Aab50093 GB1 prote
Adc64559 Trichodes
Adq07360 Trichodes
Abb71088 Drosophil
Abu44608 Protein e
Aam69065 Phototrab
Adb04473 Enterocin
Adf04643 Bacterial
Adc95053 E. faeciu
Abu42656 Protein e
Abp38314 Staphyloc
Ado84851 S epiderm
Aas04803 Staphyloc
Aay84337 Fusarium
Aab14872 Fusarium
Aau77580 F. javani
Aau77424 Fusarium
Add46266 Rat Prote
Ade62845 Rat Prote

26 59 19.4 753 4 AAB50093
27 57.5 18.9 461 7 ADC64559
28 57.5 18.9 461 8 ADQ07360
29 57.5 18.9 643 4 ABB71088
30 57 18.8 581 6 ABU44608
31 57 18.8 1565 6 AAM69065
32 56.5 18.6 45 5 ADB04473
33 56.5 18.6 242 7 ADF04643
34 56 18.4 66 7 ADC95053
35 56 18.4 6641 6 ABU42656
36 56 18.4 10182 5 ABP38314
37 56 18.4 10203 8 ADO84851
38 55.5 18.3 244 2 AAS04803
39 55.5 18.3 244 3 AAY84337
40 55.5 18.3 244 3 AAB14872
41 55.5 18.3 244 5 AAU77580
42 55.5 18.3 244 5 AAU77424
43 55.5 18.3 1382 7 ADD46266
44 55.5 18.3 1382 7 ADE62845
45 55.5 18.3 1382 7 ADE62845

Best Local Similarity 27.7%; Pred. No. 43;
Matches 13; Conservative 8; Mismatches 17; Indels 9; Gaps 2;

Qy 9 VHCTKNS-----LKGKVR-LKNMKYDQNTTYMGRLODILLGWAT 46
Db 26 IECMSNEHSSKEEMWMLWDEVKQAKQLSWRNLLYNGDFEDVSNGWKT 72

RESULT 4

AAW11633
ID AAW11633 standard; protein; 42 AA.

AC AAW11633;

XX 25-MAR-2003 (revised)

DT 05-JAN-1998 (first entry)

XX Lactobacillus plantarum bacteriocin.

XX bacteriocin; Lactobacillus plantarum BN85; DSM 9296; soft cheese;
KW inhibitor; growth; ripening; washed-rind; Munster; L. monocytogenes;
KW L. innocua; L. seeligeri; L. ivanovii; Listeria.

OS Lactobacillus plantarum.

XX Key Location/Qualifiers

FF Disulfide-bond 9..24

FT Modified-site 42

PT /note= "Cys-NH2"

XX DE29616369-UL.

XX 23-JAN-1997.

XX 20-SEP-1996; 96DE-02016369.

XX 20-SEP-1996; 96DE-02016369.

XX (FROM-) FROMAGERIE SIFPERT FRERES SA.

XX WPI; 1997-088733/09.

XX Lactobacillus strain that inhibits Listeria - for use in cheese prodn.

XX Example 4; Page 32; 53pp; German.

XX This is bacteriocin from a Lactobacillus strain [L. plantarum BN85 (DSM 9296)], which is derived from soft cheese. A partial fragment derived by Edman degradation (AAW11632) is capable of inhibiting growth of Listeria bacteria in cheese at the start of ripening. The new Lactobacillus strain is useful for prodn. of Listeria-free cheese, pref. washed-rind cheese, esp. Munster cheese. Strain DSM 9296 is added to the cheese during ripening, e.g. by spraying the cheese with a soln. contg. 104-106 cfu/ml of Lactobacillus. DSM 9296 is capable of inhibiting Listeria monocytogenes 1/2a, 1/2b, 1/2c, 3a, 3b, 4b, 4d and v7; Listeria innocua 6a and 6b; Listeria seeligeri; and Listeria ivanovii. (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 42 AA;

XX Query Match 20.4%; Score 62; DB 2; Length 42;
Best Local Similarity 39.1%; Pred. No. 2.2;
Matches 18; Conservative 6; Mismatches 8; Indels 14; Gaps 4;

Qy 3 YYGTNGVHCTKNSL-WGKVR-LKNMKYDQNTTYMGRLODILLGWATG 47
Db 2 YYG-NGVTCGKHGSDWGKA-----TTCI--INNAMAWATG 34

RESULT 5

AAW66446
ID AAW66446 standard; peptide; 37 AA.

AC AAW66446;

XX 12-JAN-1999 (first entry)

XX Cationic peptide leukocin A-val 187.

XX Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW bacterial infection; tolerance; antibacterial; microorganism; bacteria;
KW fungus; parasite; virus.

OS Leuconostoc gelidum.

XX WO9840401-A2.

XX 17-SEP-1998.

XX 10-MAR-1998; 98WO-CA000190.

XX 10-MAR-1997; 97US-0040649P.

XX 20-AUG-1997; 97US-00915314.

XX 26-SEP-1997; 97US-0060099P.

XX 25-FEB-1998; 98US-00030619.

XX (MCR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, West MHP, Mcnicol PJ;

XX WPI; 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing activity of antibiotic or overcoming tolerance, acquired resistance or inherent resistance of microorganisms.

XX Disclosure; Page 11; 105pp; English.

XX AAW66393 to AAW66469 represent native cationic peptides from the present invention. The present invention describes compositions and methods for treating infection, especially bacterial infections. The compositions and methods use cationic peptides in combination with an antibiotic agent which are then administered to a patient to enhance the activity of the antibiotic agent, to overcome: (a) tolerance; (b) acquired resistance; and (c) inherent resistance. The combinations of antibiotics and cationic peptides can provide synergistic activity against a microorganism that is tolerant, inherently resistant, or has acquired resistance to an antibiotic agent. They can be used for killing e.g. bacteria, fungi, parasites and viruses

XX Sequence 37 AA;

XX Query Match 19.7%; Score 60; DB 2; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.5;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
Db 2 YYG-NGVHCTKSGCVNMGEAFSAGVHRLAN 31

RESULT 6

AAV91745

ID AAV91745 standard; peptide; 37 AA.

XX AAV91745;

XX 06-JUN-2000 (first entry)

XX Cationic peptide Leukocin A-val 187 amino acid sequence.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.

```

OS Unidentified.
XX
XX WO9965506-A2.
XX
XX 23-DEC-1999.
XX
XX 14-JUN-1999; 99WO-CA000552.
XX
XX 12-JUN-1998; 98US-00096541.
XX
XX (MTCR-) MICROLOGIX BIOTECH INC.
XX
XX Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHP;
XX
XX WPI; 2000-223549/19.
XX
XX Novel pharmaceutical composition containing optionally activated
XX polyoxyalkylene-modified cationic peptides, useful for treating tumors.
XX
XX Disclosure; Page 11; 94pp; English.
XX
XX This sequence represents a cationic peptide amino acid sequence, which
XX can be used in the pharmaceutical composition of the invention. The
XX invention relates to a pharmaceutical composition containing at least one
XX activated polyoxyalkylene (APO)-modified cationic peptide. The
XX modification of peptides with APO increases their activity against tumour
XX cells, including those with a multidrug resistant phenotype. The
XX pharmaceutical composition can be used to treat tumours, specifically
XX lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
XX cervix, uterus, skin, prostate, liver and colon
XX
XX Sequence 37 AA;
SQ
Query Match 19.7%; Score 60; DB 3; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.5;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
QY 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
DB 2 YYG-NGVHCTKSGSVNMGFAFSAGVHRLAN 31
RESULT 7
-AAR56481
ID AAR56481 standard; protein; 300 AA.
XX
XX AAR56481;
XX
XX 25-MAR-2003 (revised)
XX 22-FEB-1995 (first entry)
XX
XX CD38.
XX
XX CD38; therapeutic; diagnostic.
XX
XX Homo sapiens.
XX
XX WO9417184-A1.
XX
XX 04-AUG-1994.
XX
XX 27-JAN-1994; 94WO-US000517.
XX
XX 29-JAN-1993; 93US-00010905.
XX
XX (SCHE ) SCHERING CORP.
XX (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
XX Parkhouse RME, Santos-Argumedo L, Grimaldi JC, Bazan JF, Heath A;
XX Howard MC, Goodnow CC;
XX
XX WPI; 1994-264098/32.
XX
XX N-PSDB; AAR70644.
XX
XX Modulation of responses of lymphocytes - using antibody to CD38, a
XX soluble CD38 fragment or a modulator of ADP-ribosyl cyclase or cyclic ADP
XX -ribosyl hydrolase.
XX
XX Disclosure; Page 46; 54pp; English.
XX
XX The protein is a human CD38 molecule which is used to modulate the
XX physiological response of a lymphocyte, which is useful for the
XX stimulation/inhibition of lymphocyte growth or differentiation,
XX particularly for the establishment of antigen tolerance. (Updated on 25-
XX MAR-2003 to correct PN field.)
XX
XX Sequence 300 AA;
SQ
Query Match 19.7%; Score 60; DB 2; Length 300;
Best Local Similarity 35.0%; Pred. No. 43;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
QY 5 GTNGVHCTKNSLWGVRLKMKYDQNTTYNGRLQDILLGW 44
DB 113 GTQVPCNKILLWSRIKDLAHQFTQVQORDMFTLEDLLGY 152
RESULT 8
AAW37928
ID AAW37928 standard; protein; 300 AA.
XX
XX AAW37928;
XX
XX 09-SEP-1998 (first entry)
XX
XX Amino acid sequence of CD38.
XX
XX CD38; CD38 peptide fragment; anti-CD38 autoantibody; autoimmune disease;
XX insulin dependent diabetes mellitus; type II diabetes.
XX
XX Mammalia.
XX
XX WO9816245-A1.
XX
XX 23-APR-1998.
XX
XX 11-APR-1997; 97WO-JP001259.
XX
XX 15-OCT-1996; 96JP-00272537.
XX
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Taminato T;
XX
XX WPI; 1998-251054/22.
XX
XX N-PSDB; AAV29155.
XX
XX Assaying anti-CD38 auto-antibody - useful for detecting auto-immune
XX disease, e.g. type II diabetes.
XX
XX Disclosure; Page 8-10; 17pp; Japanese.
XX
XX This is the amino acid sequence of the CD38 protein, of which a fragment
XX is used in the method of the invention to create the anti-CD38
XX autoantibody. The method is used for detecting autoimmune disease, e.g.
XX insulin dependent diabetes mellitus or type II diabetes
XX
XX Sequence 300 AA;
SQ
Query Match 19.7%; Score 60; DB 2; Length 300;
Best Local Similarity 35.0%; Pred. No. 43;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
QY 5 GTNGVHCTKNSLWGVRLKMKYDQNTTYNGRLQDILLGW 44
DB 113 GTQVPCNKILLWSRIKDLAHQFTQVQORDMFTLEDLLGY 152

```

PR	13-OCT-2000; 2000US-00687576.
PR	08-DEC-2000; 2000US-00732288.
PR	08-DEC-2000; 2000US-00733742.
PR	24-JAN-2001; 2001US-0263957P.
PR	16-MAR-2001; 2001US-0276791P.
PR	16-MAR-2001; 2001US-0276888P.
PR	06-APR-2001; 2001US-0281922P.
PR	24-APR-2001; 2001US-0286214P.
PR	30-APR-2001; 2001US-00847046.
PR	04-MAY-2001; 2001US-0288589P.
XX	(EOSB-) EOS BIOTECHNOLOGY INC.
PA	Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX	WPI; 2002-471335/50.
DR	N-PSDB; ABK92132.
DR	
XX	Detecting a prostate cancer-associated transcript in a cell in a patient.
PT	useful for diagnosing prostate cancer (PC) or screening modulators of PC.
PT	by determining if prostate cancer-associated genes are expressed in a
PT	prostate tissue.
XX	
XX	Claim 27; Page 314; 436pp; English.
PS	
XX	The present invention relates to methods of detecting a prostate cancer-
CC	associated transcript in a cell from a patient. The method comprises
CC	contacting a biological sample from the patient with prostate cancer-
CC	associated polynucleotides (designated PC genes) that selectively
CC	hybridise to a sequence that is at least 80% identical to them. The
CC	prostate cancer-associated polynucleotide sequences are differentially
CC	expressed in prostate tumour tissue or in prostate cancer and are derived
CC	from the tissues of various organisms such as humans or other mammals
CC	(e.g. mice, sheep and dogs). The methods of the invention are useful for
CC	diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC	associated genes are useful for diagnosing or treating prostate cancer,
CC	as well as for identifying modulators of prostate cancer or agents that
CC	inhibit prostate cancer. The nucleic acid sequences are particularly
CC	useful in gene therapy, as a vaccine or in antisense applications.
CC	ABG61800-ABG61944 represent prostate cancer-associated proteins
XX	
XX	Sequence 300 AA;
SQ	
	Query Match 19.7%; Score 60; DB 5; Length 300;
	Best Local Similarity 35.0%; Pred. No. 43;
	Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Oy	5 GTGGVHCTKSLNGKVRLKMKYDQNTTYMGRLODILLGW 44 :: : :
Db	113 GTQTVCNKILLSRKINDLAHQFTQVQRDMFTLEDTLGY 152
RESULT	
RESULT 11	
ABU09711	
ID	ABU09711 standard; protein; 300 AA.
XX	
AC	ABU09711;
XX	
DT	03-JUL-2003 (first entry)
XX	
DE	CD38 protein.
XX	
KW	CD38; diabetic onset; diabetes; lymphocyte surface marker.
XX	
OS	Hominidae.
XX	
FN	US2003027134-A1.
XX	
PD	06-FEB-2003.
XX	
Pf	31-JAN-2001; 2001US-00773307.
XX	
PR	31-JAN-2001; 2001US-00773307.

```

XX (EGAS/) EGASHIRA T.
PA (NAGA/) NAGANO M.
PA (SAGE/) SAGEHASHI Y.
PA (MATS/) MATSUI K.
PA (HATT/) HATTORI H.
PA (KANA/) KANATSUKA A.
PA (TAKA/) TAKASAWA S.
PA (OKAM/) OKAMOTO H.
XX
XX Egashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;
PI Kanatsuka A, Takasawa S, Okamoto H;
XX
XX WPI; 2003-417248/39.
DR N-PSDB; ACA60951.
XX
XX Detecting risk factor for onset of diabetes in an individual, involves
XX detecting genetic abnormality of the gene CD38.
XX
XX Disclosure; Page 10; 29pp; English.
XX
XX The invention describes a method of detecting a risk factor for diabetic
XX onset in an individual. The method involves detecting genetic abnormality
XX of the gene CD38 which is a human lymphocyte surface marker and is a risk
XX gene for onset of diabetes. The sites of abnormality in CD38 gene include
XX a site encoding Arg at residue 140, a site encoding Ser at residue 264,
XX or guanine at nucleotide position -28 in intron 7. This is the amino acid
XX sequence of wild type CD38 protein
XX
XX Sequence 300 AA;
SQ
Query Match 19.7%; Score 60; DB 6; Length 300;
Best Local Similarity 35.0%; Pred. No. 43;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
QY 5 GTNGVHCTKNSLWKGKVRLLKNNKYDQNTTYMGRLQDILLGW 44
DB 113 GTQVPCNKILLWSRIKDLAHQFTQVQDRMFTLEDTLGY 152
XX
XX RESULT 12
XX ABU09714
XX ID ABU09714 standard; protein; 300 AA.
XX AC ABU09714;
XX
XX 03-JUL-2003 (first entry)
XX
XX CD38 protein S264L mutant.
XX
XX CD38; diabetic onset; diabetes; lymphocyte surface marker; mutant;
XX mutein.
XX
XX Hominidae.
XX
XX OS
XX
XX Key Location/Qualifiers
XX FT Misc-difference 264 /note= "Wild type Ser substituted by Leu"
XX
XX US2003027134-A1.
XX
XX 06-FEB-2003.
XX
XX 31-JAN-2001; 2001US-00773307.
XX
XX 31-JAN-2001; 2001US-00773307.
XX
XX (EGAS/) EGASHIRA T.
XX (NAGA/) NAGANO M.
XX (SAGE/) SAGEHASHI Y.
XX (MATS/) MATSUI K.
XX (HATT/) HATTORI H.
XX (KANA/) KANATSUKA A.
XX
XX (TAKA/) TAKASAWA S.
XX (OKAM/) OKAMOTO H.
XX
XX Egashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;
PI Kanatsuka A, Takasawa S, Okamoto H;
XX
XX WPI; 2003-417248/39.
DR N-PSDB; ACA60951.
XX
XX Detecting risk factor for onset of diabetes in an individual, involves
XX detecting genetic abnormality of the gene CD38.
XX
XX Disclosure; Page 10; 29pp; English.
XX
XX The invention describes a method of detecting a risk factor for diabetic
XX onset in an individual. The method involves detecting genetic abnormality
XX of the gene CD38 which is a human lymphocyte surface marker and is a risk
XX gene for onset of diabetes. The sites of abnormality in CD38 gene include
XX a site encoding Arg at residue 140, a site encoding Ser at residue 264,
XX or guanine at nucleotide position -28 in intron 7. This is the amino acid
XX sequence of wild type CD38 protein
XX
XX Sequence 300 AA;
SQ
Query Match 19.7%; Score 60; DB 6; Length 300;
Best Local Similarity 35.0%; Pred. No. 43;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
QY 5 GTNGVHCTKNSLWKGKVRLLKNNKYDQNTTYMGRLQDILLGW 44
DB 113 GTQVPCNKILLWSRIKDLAHQFTQVQDRMFTLEDTLGY 152
XX
XX RESULT 13
XX ADD18904
XX ID ADD18904 standard; protein; 300 AA.
XX AC ADD18904;
XX
XX 15-JAN-2004 (first entry)
XX
XX Human disease related protein SeqID393.
XX
XX human; disease state; cytostatic; antiinflammatory; ophthalmological;
XX antiarteriosclerotic; vulnery; gene therapy;
XX hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
XX inflammation; erythropoiesis; glycolysis; gluconeogenesis;
XX glucose transportation; catecholamine synthesis; iron transport;
XX nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
XX retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
XX inflammatory condition; wound healing.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO2003018621-A2.
XX
XX 06-MAR-2003.
XX
XX 23-AUG-2002; 2002WO-GB003892.
XX
XX 23-AUG-2001; 2001GB-00020558.
XX
XX 05-OCT-2001; 2001GB-00024037.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
XX WPI; 2003-290046/28.
XX
XX N-PSDB; ADD18905.
XX
XX New substantially purified polypeptide, useful for diagnosing or treating
XX a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
XX injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
XX

```

PT wound healing.

PS Claim 25; SEQ ID NO 393; 424pp; English.

XX

CC This invention relates to novel human genes and gene product which are implicated in certain disease states. Compounds which modulate the proteins of the invention may have cytotostatic, antiinflammatory, ophthalmological, antiarteriosclerotic or vulnary activities. The sequences of the invention may be useful for gene therapy. The invention may be useful for diagnosing or treating a hypoxia-regulated condition, such as tumorigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis, or the biological response to hypoxia conditions including processes such as glycolysis, gluconeogenesis, glucose transportation, catecholamine synthesis, iron transport or nitric oxide synthesis. The disease includes cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis, inflammatory conditions or wound healing. The present sequence is that of a disease related protein of the invention.

XX

CC Sequence 300 AA;

XX

Query Match 19.7%; Score 60; DB 7; Length 300;
Best Local Similarity 35.0%; Pred. No. 43;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWGKVRLKNMKYDQNTTYMGRLODILLGW 44
||| ||| :: :| |::| |::|
Db 113 GTQTVPCKILLSRIKDLAHOFTQVRDMFTLEDTLLGY 152

RESULT 14

ADD45262

ID ADD45262 standard; protein; 300 AA.

XX

AC ADD45262;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human Protein P28907, SEQ ID NO 10695.

XX

KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX

QS Homo sapiens.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

Pf 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

(GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

FI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

DR GENBANK; P28907.

XX

PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

FT

PS Claim 1; Page; 1017pp; English.

XX

XX The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2005, 17:27:38 ; Search time 22 Seconds
(without alignments)
183.230 Million cell updates/sec

Title: US-10-644-927-1
Perfect score: 304
Sequence: 1 KTYGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69.5	22.9	420	4	US-09-328-352-5261
2	64	21.1	61	4	US-09-068-507C-5
3	62	20.4	85	4	US-09-270-767-44549
4	60	19.7	36	4	US-08-924-629C-69
5	60	19.7	37	4	US-09-030-619-209
6	60	19.7	61	4	US-08-924-629C-68
7	59.5	19.6	577	4	US-09-248-796A-18807
8	59	19.4	62	4	US-08-924-629C-71
9	56.5	18.6	242	4	US-09-543-681A-4928
10	56	18.4	66	4	US-09-107-532A-4680
11	56	18.4	10182	3	US-09-114-001C-3159
12	55.5	18.3	243	3	US-09-216-295-13
13	55.5	18.3	244	4	US-09-632-570-13
14	55.5	18.3	244	4	US-09-632-575-43
15	55	18.1	41	4	US-08-924-629C-70
16	55	18.1	177	4	US-09-248-796A-18521
17	54.5	17.9	494	3	US-08-378-313-23
18	54.5	17.9	494	3	US-08-378-313-29
19	54.5	17.9	641	4	US-09-071-035-456
20	54.5	17.9	1313	4	US-09-071-035-450
21	54.5	17.9	1313	4	US-09-071-035-454
22	54	17.8	697	4	US-09-489-039A-7485
23	53.5	17.6	56	4	US-09-834-309-3
24	53.5	17.6	74	4	US-09-107-532A-5289
25	53.5	17.6	134	4	US-09-834-309-4
26	53.5	17.6	1033	4	US-09-834-309-1
27	53	17.4	233	1	US-08-032-848C-12

28	53	17.4	263	3	US-09-216-295-24	Sequence 24, Appl
29	53	17.4	264	4	US-09-632-570-24	Sequence 24, Appl
30	53	17.4	264	4	US-09-632-575-54	Sequence 54, Appl
31	53	17.4	384	3	US-09-311-170-2	Sequence 2, Appl
32	53	17.4	929	4	US-09-134-000C-6424	Sequence 6424, Ap
33	52.5	17.3	371	4	US-09-248-796A-17748	Sequence 17748, A
34	52.5	17.3	418	4	US-09-610-104C-12	Sequence 12, Appl
35	52.5	17.3	911	3	US-09-356-952-6	Sequence 6, Appl
36	52.5	17.3	1176	1	US-08-257-999-2	Sequence 2, Appl
37	52	17.1	219	4	US-09-540-236-2633	Sequence 2633, Ap
38	52	17.1	221	4	US-09-270-767-33753	Sequence 33753, A
39	52	17.1	249	4	US-09-282-218A-17	Sequence 17, Appl
40	52	17.1	459	2	US-08-810-655A-2	Sequence 2, Appl
41	52	17.1	492	4	US-09-540-236-3061	Sequence 3061, Ap
42	51.5	16.9	139	4	US-09-902-540-13820	Sequence 13820, A
43	51.5	16.9	360	4	US-09-902-540-11642	Sequence 11642, A
44	51.5	16.9	510	4	US-09-727-628-2	Sequence 2, Appl
45	51	16.8	59	4	US-08-924-628C-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-09-328-352-5261
; Sequence 5261, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5261
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5261

Query Match 22.9%; Score 69.5; DB 4; Length 420;
Best Local Similarity 38.5%; Pred. No. 0.41;
Matches 20; Conservative 8; Mismatches 13; Indels 11; Gaps 4;

Qy 1 KTYGTN--GVHCTKNSLWGVKVLKXMKY--DQNTYMGRLQDI---LLGW 44
Db 361 ESYVTNPKSGKSHNKANIWIK---NNTPYSGDRDDTYTTLRLDITSPILGW 409

RESULT 2
US-09-068-507C-5
; Sequence 5, Application US/09068507C
; Patent No. 6790951
; GENERAL INFORMATION:
; APPLICANT: EUSINK, VINCENT et al.
; TITLE OF INVENTION: EXPRESSION SYSTEM IN MICROORGANISM AND ITS USE FOR EXPRESSING
; FILE REFERENCE: 1380-0122P
; CURRENT APPLICATION NUMBER: US/09/068,507C
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Lactobacillus sake
US-09-068-507C-5

Query Match 21.1%; Score 64; DB 4; Length 61;
Best Local Similarity 36.7%; Pred. No. 0.2;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

Qy	3	Y Y G T N G V H C T K N S L --- W G K V R L K N M K Y D Q N T T Y M G R L - Q D I L L G W A T G	47
		:	
		: :	
		: :	
Db	20	Y Y G - N G V H C G K H S C T V D W G ----- T A I G N I G N N A A A N W A T G	54

RESULT 3

```

US-09-270-767-44549
; Sequence 44549, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44549
; ~ LENGTH: 85
; TYPE: PR1
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid

```

RESULT 4

```

US-08-924-629C-69
; Sequence 69, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aislin
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Mesenteriocin Y105
US-08-924-629C-69

```

RESULT 5

```

US-09-030-619-209
; Sequence No. 6503881
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND
; TITLE OF INVENTION: INFECTIONS USING
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version:
; SEQ ID NO 209
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidum
US-09-030-619-209

```

RESULT 6

```

US-08-924-629C-68
; Sequence 68, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Rodney W.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aislin
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082BacteriaIocins, Transport an
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Leucocin A
; US-08-924-629C-68

```

RESULT 7

US-09-248-796A-18807

; Sequence 18807, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; PRIOR FILING DATE: 1999-02-12

; CURRENT APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 18807

; LENGTH: 577

; TYPE: PRT

; ORGANISM: Candida albicans

; US-09-248-796A-18807

Query Match 19.6%; Score 59.5; DB 4; Length 577;

Best Local Similarity 30.6%; Pred. No. 14;

Matches 15; Conservative 7; Mismatches 12; Indels 15; Gaps 2;

.QY 6 TNGVHCTKNSLWGLVKLNKMYDQNTTMYMGRLODILLGWATGAFGKTFH 54

Db 261 TTGVCQDESTLWVKL-IPNLKHLNN-----QTSAGKPFH 294

RESULT 8

US-08-924-629C-71

; Sequence 71, Application US/08924629C

; Patent No. 6403082

; GENERAL INFORMATION:

; APPLICANT: Stiles, Michael E.

; APPLICANT: Vederas, John C.

; APPLICANT: van Belkum, Marius J.

; APPLICANT: Worobo, Randy W.

; APPLICANT: Worobo, Rodney J.

; APPLICANT: Greer, G. Gordon

; APPLICANT: McMullen, Lynn M.

; APPLICANT: Leisner, Jorgen J.

; APPLICANT: Poon, Alison

; APPLICANT: Franz, Charles M.A.P.

; TITLE OF INVENTION: No. 6403082elbacteriocins, Transport and Vector System and Method

; FILE REFERENCE: 660.0005US

; CURRENT APPLICATION NUMBER: US/08/924,629C

; CURRENT FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: US 60/026,257

; PRIOR FILING DATE: 1996-09-05

; NUMBER OF SEQ ID NOS: 80

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 71

; LENGTH: 62

; TYPE: PRT

; ORGANISM: Pediocin PA1

; US-08-924-629C-71

Query Match

Best Local Similarity 19.4%; Score 59; DB 4; Length 62;

Matches 18; Conservative 5; Mismatches 9; Indels 16; Gaps 4;

.QY 3 YGTNGVHCTKNSL---WGLVKLNKMYDQNTTMYMGRLODILLGWATG 47

Db 20 YYG-NGVTCGKSCSDWGKA-----TTCTI---INNGAWATG 54

RESULT 9

US-09-543-681A-4928

; Sequence 4928, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 4928

; LENGTH: 242

; TYPE: PRT

; ORGANISM: Proteus mirabilis

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (18)

; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

; US-09-543-681A-4928

Query Match 18.6%; Score 56.5; DB 4; Length 242;

Best Local Similarity 28.1%; Pred. No. 12;

Matches 18; Conservative 8; Mismatches 21; Indels 17; Gaps 3;

.QY 4 YGTNGVHCTKNSLWGLVKLNK-----MKYDQNTTMYMGRLODILLGWATGAF 49

Db 106 YKTSSTDSAKISVFLVNAKNTEQIGYIIDVIDIKVDNKRAYLGRVADKI--W-TGSS 162

.QY 50 GKTF 53

Db 163 ATIF 166

RESULT 10

US-09-107-532A-4680

; Sequence 4680, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 4680:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 66 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

Handwritten signature/initials

; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aleson
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Sakacin P
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (33)..(33)
; OTHER INFORMATION: any amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (37)..(37)
; OTHER INFORMATION: any amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (40)..(40)
; OTHER INFORMATION: any amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (41)..(41)
; OTHER INFORMATION: any amino acid
; US-08-924-629C-70

Query Match 18.1%; Score 55; DB 4; Length 41;
Best Local Similarity 63.2%; Pred. No. 2.1;
Matches 12; Conservative 1; Mismatches 2; Indels 4; Gaps 2;

QY 3 YGTGNGVHCTKNSL---WG 18
||| ||||| |:
Db 2 YYG-NGVHCGKHSCTVDWG 19

Search completed: May 16, 2005, 17:35:11
Job time : 23 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2005, 17:34:06 ; Search time 134 Seconds
(without alignments)
134.618 Million cell updates/sec

Title: US-10-644-927-1

Perfect score: 304

Sequence: 1 KTYGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1432185 seqs, 334051727 residues

*Total number of hits satisfying chosen parameters: 1432185

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

*Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	19.7	36	10 US-09-883-343A-69	Sequence 69, Appl
2	60	19.7	37	9 US-09-030-619-209	Sequence 209, App
3	60	19.7	37	9 US-09-912-609-97	Sequence 97, Appl
4	60	19.7	37	15 US-10-277-232-209	Sequence 209, App
5	60	19.7	37	15 US-10-277-233-209	Sequence 209, App
6	60	19.7	61	10 US-09-883-343A-68	Sequence 68, Appl
7	60	19.7	300	9 US-09-982-616-9	Sequence 9, Appli
8	60	19.7	300	10 US-09-773-307B-2	Sequence 2, Appli
9	60	19.7	300	15 US-10-295-027-1348	Sequence 1348, Ap
10	60	19.7	300	15 US-10-454-238-2	Sequence 2, Appli
11	60	19.7	300	16 US-10-322-696-162	Sequence 162, App
12	60	19.7	300	17 US-10-753-267-34	Sequence 34, Appl
13	59	19.4	62	10 US-09-883-343A-71	Sequence 71, Appl

Sequence 6141, Ap
Sequence 280304,
Sequence 213009,
Sequence 10, Appl
Sequence 10, Appl
Sequence 197752,
Sequence 72532, A
Sequence 225665,
Sequence 11, Appl
Sequence 70580, A
Sequence 23, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 70314, A
Sequence 60740, A
Sequence 70, Appl
Sequence 5582, Ap
Sequence 5591, A
Sequence 103806,
Sequence 456, App
Sequence 456, App
Sequence 450, App
Sequence 454, App
Sequence 450, App
Sequence 56879, A
Sequence 454, App
Sequence 454, App
Sequence 10936, A
Sequence 732, App
Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-883-343A-69
; Sequence 69, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aislin
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: NO. US20030039632A1eBacteriocins, Transport and Vector System and
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/09/883,343A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US/08/924,629
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Mesenteriocin Y105
US-09-883-343A-69

Query Match 19.7%; Score 60; DB 10; Length 36;
Best Local Similarity 48.4%; Pred.No. 2.3;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
QY 3 YYGTNGVHCTKNSL----WGKV-----RLKN 24


```

; APPLICANT: BML, INC.
; TITLE OF INVENTION: Method of Detecting Risk Factor for Onset of Diabetes
; FILE REFERENCE: PEM37
; CURRENT APPLICATION NUMBER: US/09/773,307B
; CURRENT FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Hominidae
US-09-773-307B-2

Query Match      19.7%; Score 60; DB 10; Length 300;
Best Local Similarity 35.0%; Pred. No. 25;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy   5 GTGVTCTKNSLKGKVLKMKYDKDNTTYNGRLQDLILGW 44
Db   113 GTQTVCNKILLWSRIKDLAHQFTQVORDMFTLEDTLLGY 152

RESULT 9
US-10-295-027-1348
; Sequence 1348, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1348
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1348

Query Match      19.7%; Score 60; DB 15; Length 300;
Best Local Similarity 35.0%; Pred. No. 25;

```

```
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy 5 GTNGVHCTKNSLWKGKVLKMKYDQNTYMGRLQDILLGW 44
Db 113 GTQTPCNKILLWSRIKDLAHQFTQVQORDMFTLEDTLGGY 152

RESULT 10
US-10-454-238-2
; Sequence 2, Application US/10454238
; Publication No. US20040081981A1
; GENERAL INFORMATION:
; APPLICANT: Eguashiro et al.
; TITLE OF INVENTION: Method of Detecting Risk Factor for Onset of Diabetes
; FILE REFERENCE: WSHING.001DV1
; CURRENT APPLICATION NUMBER: US/10/454,238
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: US 09/773,307
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Hominidae
US-10-454-238-2

Query Match 19.7%; Score 60; DB 15; Length 300;
Best Local Similarity 35.0%; Pred. No. 25;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy 5 GTNGVHCTKNSLWKGKVLKMKYDQNTYMGRLQDILLGW 44
Db 113 GTQTPCNKILLWSRIKDLAHQFTQVQORDMFTLEDTLGGY 152

RESULT 11
US-10-322-696-162
; Sequence 162, Application US/10322696
; Publication No. US20040166490A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Malandro, Marc
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 529452001200
; CURRENT APPLICATION NUMBER: US/10/322,696
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-322-696-162

Query Match 19.7%; Score 60; DB 16; Length 300;
Best Local Similarity 35.0%; Pred. No. 25;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy 5 GTNGVHCTKNSLWKGKVLKMKYDQNTYMGRLQDILLGW 44
Db 113 GTQTPCNKILLWSRIKDLAHQFTQVQORDMFTLEDTLGGY 152

RESULT 12
US-10-753-267-34
; Sequence 34, Application US/10753267
; Publication No. US20050037946A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Stagliano, Nancy E.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
```

```
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Donoghue, Mary A.
; APPLICANT: Rodriguez-Way, Amelie
; APPLICANT: Tomlinson, James E.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 1722, 10280, 59917, 85553,
; TITLE OF INVENTION: 10653, 9235, 21668, 17794, 2210, 6169, 10102, 21061, 17662,
; TITLE OF INVENTION: 1468, 12282, 6350, 9035, 1820, 23652, 7301, 8925, 8701,
; TITLE OF INVENTION: 3533, 9462, 9123, 12788, 17729, 55552, 1261, 21476, 33770,
; TITLE OF INVENTION: 9380, 2569654, 33556, 53656, 44143, 32612, 10671, 261,
; TITLE OF INVENTION: 44570, 41922, 2552, 2417, 19319, 43969, 8921, 8993, 955,
; TITLE OF INVENTION: 32345, 966, 1920, 17318, 1510, 14180, 26005, 554, 16408,
; TITLE OF INVENTION: 42028, 112091, 13886, 13942, 1673, 54946 OR 2419
; FILE REFERENCE: MPI03-003PIRNOMNIM
; CURRENT APPLICATION NUMBER: US/10/753,267
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 60/439,683
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/445,216
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,036
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/454,189
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/457,541
; PRIOR FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/466,411
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/469,041
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/477,414
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/478,560
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/489,772
; PRIOR FILING DATE: 2003-07-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-753-267-34

Query Match 19.7%; Score 60; DB 17; Length 300;
Best Local Similarity 35.0%; Pred. No. 25;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy 5 GTNGVHCTKNSLWKGKVLKMKYDQNTYMGRLQDILLGW 44
Db 113 GTQTPCNKILLWSRIKDLAHQFTQVQORDMFTLEDTLGGY 152

RESULT 13
US-09-883-343A-71
; Sequence 71, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Rodney W.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Joergen J.
; APPLICANT: Poon, Alsaion
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. US20030039632A1e1Bacteriocins, Transport and Vector System an
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/09/883,343A
; CURRENT FILING DATE: 2001-06-19
```

; PRIOR APPLICATION NUMBER: US/08/924,629
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Pedicocin PA1
US-09-883-343A-71

Query Match 19.4%; Score 59; DB 10; Length 62;
Best Local Similarity 37.5%; Pred. No. 5.8;
Matches 18; Conservative 5; Mismatches 9; Indels 16; Gaps 4;

QY 3 YGCTNGVHCTKNSL---WGKRLKNMKYDQNTTYMGLQDILLGWATG 47
Db 20 YYG-NGVTCGKHSVDWGKA-----TTCTI--INNAMAWATG 54

RESULT 14
US-10-369-493-6141
; Sequence 6141, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6141
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6141

*Query Match 19.4%; Score 59; DB 15; Length 399;
Best Local Similarity 35.8%; Pred. No. 47;
Matches 19; Conservative 4; Mismatches 22; Indels 8; Gaps 3;

QY 3 YGCTNGVHCTKNSLWGKVKRLKNMKYDQNTTYMGLQDILLG--WATGAFGKTF 53
Db 75 YYTDSGSCSN----GRFQLKNFILQOPWEYMH--SDVTAGKVLGEGAGKVF 121

RESULT 15
US-10-424-599-280304
; Sequence 280304, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 280304
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Glycine max

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_95137C.1.pep
US-10-424-599-280304

Query Match 19.2%; Score 58.5; DB 15; Length 95;
Best Local Similarity 35.3%; Pred. No. 11;
Matches 18; Conservative 5; Mismatches 19; Indels 9; Gaps 3;

QY 1 KTYVGTNGV---HCTKNLSLWGKVKRLKNMKYDQNTTYMGLQDILL-GWATG 47
Db 38 ETFLGDNGVAPFOITGNQTWSVFTTLK-----TRMGGLGPLIKGWP TG 83

Search completed: May 16, 2005, 17:46:42
Job time : 135 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2005, 17:25:23 ; Search time 39 Seconds
(without alignments)
133.223 Million cell updates/sec

Title: US-10-644-927-1
Perfect score: 304
Sequence: 1 KTYVGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTFH 54
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65.5	21.5	471	2	S11899
2	65.5	21.5	471	2	A99180
3	64	21.1	61	1	S38508
4	63	20.7	590	2	AF2555
5	62.5	20.6	446	2	A43995
6	60	19.7	61	2	A41657
7	60	19.7	61	2	S52208
8	60	19.7	300	2	A43521
9	59.5	19.6	286	2	F84954
10	59	19.4	62	1	A48941
11	59	19.4	399	2	T20455
12	58	19.1	3036	2	T18995
13	57.5	18.9	510	2	T50021
14	56.5	18.6	517	1	G1BPT4
15	55.5	18.3	132	2	PC2131
16	55.5	18.3	934	2	B29838
17	55.5	18.3	1332	2	F69732
18	55	18.1	201	2	AD3302
19	55	18.1	346	2	B90243
20	55	18.1	369	2	T24205
21	55	18.1	516	2	G84442
22	55	18.1	799	1	TVRTTB
23	54.5	17.9	319	2	T15463
24	54.5	17.9	332	2	D86760
25	54.5	17.9	374	2	F97257
26	54.5	17.9	383	2	B96806
27	54.5	17.9	494	1	B41141
28	54	17.8	262	1	TLBPM1
29	54	17.8	269	2	S51815

30 54 17.8 494 2 JC2382 sodium/proline sym
31 53.5 17.6 254 2 S31393 chlorophyll a/b-bi
32 53.5 17.6 324 2 G69515 transcription regu
33 53.5 17.6 462 1 YSBVTM threonine-tRNA lig
34 53.5 17.6 510 2 D84610 probable myo-inosi
35 53.5 17.6 527 2 S29530 gene 12 protein -
36 53.5 17.6 1091 1 PL0009 complement C3d/Bps
37 53.5 17.6 1166 2 H71609 hypohetical prote
38 53.5 17.6 1585 2 B69948 phage-related prot
39 53 17.4 167 2 S49628 FN81 protein - ye
40 53 17.4 202 2 C90228 hypohetical prote
41 53 17.4 249 2 B95357 probable Etfa2 ele
42 53 17.4 258 2 T13591 tail fiber adhesin
43 53 17.4 264 2 JU0328 cellulase (EC 3.2.
44 53 17.4 327 2 AI3308 glutathione transf
45 53 17.4 327 2 S54560 TOM37 protein - ye

ALIGNMENTS

RESULT 1

S11899

Glutamate-ammonia ligase (EC 6.3.1.2) - Sulfolobus solfataricus

N;Alternate names: Glutamine synthetase

C;Species: Sulfolobus solfataricus

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C;Accession: S11899

R;Sanangelantonio, A.M.; Barbarini, D.; di Pasquale, G.; Cammarano, P.; Tiboni, O. Mol. Gen. Genet. 221, 187-194, 1990

A;Title: Cloning and nucleotide sequence of an archaeobacterial glutamine synthetase gene

A;Reference number: S11899; MUID:90318316; PMID:1973523

A;Accession: S11899

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-471 <SAN>

A;Cross-references: UNIPROT:P23794; GB:X53263; NID:g296864; PIDN:CAA37353.1; PID:g296865

C;Superfamily: Glutamate-ammonia ligase

C;Keywords: ligase

Query Match 21.5%; Score 65.5; DB 2; Length 471;
Best Local Similarity 48.6%; Pred. No. 1.8;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

Qy 1 KTYVGTN--GVHCTKNSLWGVKRLKNKMYDQNTTY 33
Db 255 KPIYDNGTGMH-THLSLWTKGKKKLMYPDNDY 288

RESULT 2

A99180

hypohetical protein glnA-1 [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C;Accession: A99180

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Regan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: A99180

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-471 <KOR>

A;Cross-references: UNIPROT:P23794; GB:AE006641; NID:g13813510; PIDN:AAK40696.1; GSPDB:G13813510

C;Gene: glnA-1

C;Superfamily: glutamate-ammonia ligase

Query Match 21.5%; Score 65.5; DB 2; Length 471;
Best Local Similarity 48.6%; Pred. No. 1.8;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

Qy 1 KTYVGTN--GVHCTKNSLGKVRLLKNMKYDNTTYY 33
| | | | : | | | | | | | |
Db 255 KPIYGNDGTGMH--THLSLWTKDGKKNLMPDNDEY 288

RESULT 3

S38508
sakacin P precursor - Lactobacillus sake (strains Lb674 and LTH673)
C:Species: Lactobacillus sake
A:Variety: strains Lb 674 and LTH 673
C>Date: 25-Dec-1994 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S57911; S57915; S43689; S38508
R:Huehne, K.; Holick, A.; Axelsson, L.; Kroeckel, L.
submitted to the EMBL Data Library, February 1995
A:Description: Purification and cloning of sakacin 674, a bacteriocin from Lactobacillus
A:Reference number: S57910
A:Accession: S57911
A:Molecule type: DNA
A:Residues: 1-61 <HUE>
A:Cross-references: UNIPROT:P35618; EMBL:Z48542; NID:g695615; PIDN:CAA88428.1; PID:g695615
A:Experimental source: strain Lb674
R:Holick, A.L.; Axelsson, L.; Huehne, K.; Kroeckel, L.
FEMS Microbiol. Lett. 115, 143-150, 1994
A:Title: Purification and cloning of sakacin 674, a bacteriocin from Lactobacillus sake
A:Reference number: S57915; MUID:94186010; PMID:8138128
A:Accession: S57915
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-61 <HOL>
A:Cross-references: EMBL:Z25816; NID:g414124; PIDN:CAA81064.1; PID:g414125
A:Experimental source: strain Lb674
R:Tichaczek, P.S.; Vogel, R.F.; Hammes, W.P.
Microbiology 140, 361-367, 1994
A:Title: Cloning and sequencing of sakP encoding sakacin P, the bacteriocin produced by
A:Reference number: S43688; MUID:94236240; PMID:8180701
A:Accession: S43689
A:Molecule type: DNA
A:Residues: 1-61 <TIC>
A:Cross-references: EMBL:X75081; NID:g475950; PIDN:CAA52974.1; PID:g475952
A:Experimental source: strain LTH 673
C:Genetics:
A:Gene: sakP; sppA
C:Superfamily: Carnobacteriocin precursor
C:Keywords: antibacterial; bacteriocin
F:1-18/Domain: propeptide #status predicted <PRO>
F:13-61/Product: sakacin P #status predicted <MAT>
F:27-32/Disulfide bonds: #status predicted

* Query Match 21.1%; Score 64; DB 1; Length 61;
Best Local Similarity 36.7%; Pred. No. 0.28;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

Qy 3 YYFGNGVHCTKNSL---WGKVRLLKNMKYDNTTVMGR-L-QDILLGWATG 47
||| ||||| | | | | | | | | : | | | |
Db 20 YYG-NGVHGCRHSCTVDWG-----TAIGNIGNNAANWATG 54

RESULT 4

AF2555
hypothetical protein ali8037 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ga
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. Strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2555
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2555
A>Status: preliminary
A:Molecule type: DNA

```

RESULT 9
F84954
spermidine synthase (EC 2.5.1.16) [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: F84954
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: F84954
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: speE; BU209
C:Superfamily: spermidine synthase
C:Keywords: transferase

Query Match          19.6%; Score 59.5; DB 2; Length 286;
Best Local Similarity 42.1%; Pred. No. 6.2;
Matches 16; Conservative 4; Mismatches 9; Indels 9; Gaps 2;

Qy 2 TTY-----GTNGVHCTKNSLWG-KVRLKNMKYDQN 30
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 TTYGGVMVFANGTNTIEYRKNSLEKIQIRKNTKLDEN 262

```

RESULT 10
A48941
pediocin PA-1 precursor - *Pediococcus acidilactici* plasmid pSRQ11
C/Species: *Pediococcus acidilactici*
C/Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A48941; S21699; A47680; A41018; A41823
R/Variou: J.D.; Gonzalez, C.F.; Kunka, B.S.; Ledebøer, A.M.; Pucci, M.J.; Toonen, M.Y.;
Appl. Environ. Microbiol. 58, 2360-2367, 1992
A/Title: Cloning, expression, and nucleotide sequence of genes involved in production o
A/Reference number: A48941; PMID:92384551; PMID:1514784

Appl. Environ. Microbiol. 58, 2360-2367, 1992

A;/Title: Cloning, expression, and nucleotide sequence of genes involved in production of

A;/Reference number: A48941; MUID:92384551; PMID:1514784

A;/Accession: A48941

A;/Molecule type: DNA

A;/Residues: 1-62 <MAR>

A;/Cross-references: UNIPROT:P29430; GB:M83924; NID:g150565; PIDN:AAA25559.1; PID:g1505656

A;/Experimental source: strain PAC1.0, plasmid PSRQ11

A;/Note: sequence extracted from NCBI backbone (NCBIN:112082; NCBIPT:112083)

R;/Henderson, J.T.; Chopko, A.L.; van Wassenaar, P.D.

Arch. Biochem. Biophys. 295, 5-12, 1992

A;/Title: Purification and primary structure of pediocin PA-1 produced by *Pediococcus* ac

A;/Reference number: S21699; MUID:92246549; PMID:1575516

A;/Accession: S21699

A;/Molecule type: protein

A;/Residues: 19-62 <HEN>

R;/Nieto Lozano, J.C.; Meyer, J.N.; Sletten, K.; Pelaz, C.; Nes, I.F.

J. Gen. Microbiol. 138, 1985-1990, 1992

A;/Title: Purification and amino acid sequence of a bacteriocin produced by *Pediococcus*

A;/Reference number: A47680; MUID:93019000; PMID:1402795

A;/Accession: A47680

A;/Molecule type: protein

A;/Residues: 19-50, 'X', 52-61, 'X' <NIE>

A;/Note: sequence extracted from NCBI backbone (NCBIPT:116654)

A;/Note: the unidentified residues were suggested to be Trp and Cys

C;/Genetics:

A;/Gene: pedA

A;Gene: peda
A;Genome: plasmid
C;Superfamily: carnobacteriocin precursor
C;Keywords: antibacterial, bacteriocin
F;1-18/Domain: propeptide #status predicted <PRO>
F;19-62/Product: pediocin PA-1 #status experimental
F;27-32, 42-62/Disulfide bonds: #status experimental

Search completed: May 16, 2005, 17:34:43
Job time : 41 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2005, 17:17:02 ; Search time 167 Seconds
(without alignments)
165.582 Million cell updates/sec

Title: US-10-644-927-1

Perfect score: 304

Sequence: 1 KTYGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	64.1	81	2 Q48496	Q48496 lactobacilli
2	65.5	21.5	471	1 GLNA SULSO	P23794 sulfobolus
3	64.5	21.2	1071	2 Q75C66	Q75C66 ashbya goss
4	64	21.1	61	1 SAKP_LACSK	P35618 lactobacilli
5	63	20.7	590	2 Q8YK80	Q8YK80 anabaena sp
6	63	20.7	1582	2 Q8RH71	Q8RH71 fusobacteri
7	63	20.7	1630	2 Q8RH71	Q8RH71 fusobacteri
8	62.5	20.6	446	1 GLNA METVO	P21154 methanococc
9	62.5	20.6	558	2 Q8VW63	Q8VW63 bacillus th
10	61.5	20.2	469	2 Q9P203	Q9P203 xestia c-ni
11	61.5	20.2	547	2 Q8AJJ6	Q8AJJ6 bacteroides
12	61.5	20.2	573	2 Q7QBD6	Q7QBD6 anopheles g
13	61.5	20.2	853	2 Q9DL93	Q9DL93 human immun
14	61.5	20.2	858	2 Q9DL79	Q9DL79 human immun
15	60	19.7	61	1 LCCA LEUGE	P34034 leuconostoc
16	60	19.7	61	1 LCCA LEUGE	P34034 leuconostoc
17	60	19.7	61	1 MTCY LEUME	P38577 leuconostoc
18	60	19.7	61	2 Q79IV9	Q79IV9 leuconostoc
19	60	19.7	300	1 CD38 HUMAN	P28907 homo sapien
20	60	19.7	300	2 Q96HY4	Q96HY4 homo sapien
21	60	19.7	1143	2 Q948Y6	Q948Y6 volvox cart
22	60	19.7	1794	2 Q8RH71	Q8RH71 fusobacteri
23	60	19.7	1868	2 Q8IAL9	Q8IAL9 plasmodium
24	59.5	19.6	42	1 BAVM LACSK	P80493 lactobacilli
25	59.5	19.6	286	1 SPEE_BUCAI	P57305 buchneza ap
26	59	19.4	62	1 PPAL_PEDAC	P29430 pediococcus
27	59	19.4	62	2 Q8RLJ6	Q8RLJ6 pediococcus
28	59	19.4	62	2 Q53278	Q53278 pediococcus
29	59	19.4	62	2 Q79IK0	Q79IK0 lactobacilli
30	59	19.4	62	2 Q79IK3	Q79IK3 pediococcus
31	59	19.4	62	2 Q79IK6	Q79IK6 pediococcus

32 59 19.4 62 2 Q9EZB2 Q9ezb2 bacillus co
33 59 19.4 67 2 Q47778 Q47778 enterococcu
34 59 19.4 299 2 Q65IN6 Q65in6 bacillus li
35 59 19.4 419 2 Q17755 Q17755 caenorhabdi
36 59 19.4 498 2 Q9N502 Q9n502 caenorhabdi
37 59 19.4 860 2 Q8WTM9 Q8wtm9 caenorhabdi
38 59 19.4 2589 2 Q7YYI7 Q7yyi7 cryptospori
39 58.5 19.2 66 2 Q9Z4J1 Q9z4j1 carnobacter
40 58.5 19.2 485 2 Q94I16 Q94i16 cucurbita m
41 58.5 19.2 853 2 Q9DL89 Q9dl89 human immun
42 58.5 19.2 853 2 Q9DL91 Q9dl91 human immun
43 58 19.1 1951 2 Q7P2K8 Q7p2k8 fusobacteri
44 58 19.1 3118 2 Q17575 Q17575 caenorhabdi
45 57.5 18.9 129 2 Q8TKY8 Q8tky8 methanosarc

ALIGNMENTS

RESULT 1
Q48496 PRELIMINARY; PRT; 81 AA.
AC Q48496;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Acidocin A precursor.
DE Name-acda;
GN Lactobacillus acidophilus.
OS Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95314239; PubMed=7793908;
RA Kanatani K, Ohimura M., Sano K.;
RT Isolation and characterization of acidocin A and cloning of the
RT bacteriocin gene from Lactobacillus acidophilus.";
RL Appl. Environ. Microbiol. 61:1061-1067(1995).
DR EMBL=D57881; Genbank=U07120.1;
DR InterPro; IPRO10133; Bacteriocin_sig.
DR TIGRFAMS; TIGR01847; bacteriocin_sig; 1.
KW SIGNAL.
FT CHAIN 1 23 mature acidocin A.
FT CHAIN 24 81
SQ SEQUENCE 81 AA; 8946 MW; 89698AA296F7819D CRC64;
Query Match 64.1%; Score 195; DB 2; Length 81;
Best Local Similarity 64.6%; Pred. No. 1.1e-17;
Matches 42; Conservative 1; Mismatches 4; Indels 18; Gaps 3;

QY 1 KTYGTNGVHCTKNSLWGVKRLKMKYDQNTYMGRL-----QD--ILLGWATGAF 49
DB 24 KTYGTNGVHCTKNSLWGVKRLKMKYDQNTYMGRL-----IPGTLCKRQSLPIKQDLKILLGWATGAF 76
QY 50 GKTFH 54
DB 77 GKTFH 81

RESULT 2

GLNA_SULSO STANDARD; PRT; 471 AA.
ID GLNA_SULSO
AC P23794;
DT 01-NOV-1991 (Rel. 20, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).
GN Name-glNA; Synonyms=glNA-1; OrderedLocustNames=SSO0366;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/> or send an email to license@ebi.ac.uk).

```

CC      EMBL; X75081; CAA52974.1; -
CC      EMBL; Z25816; CAA81064.1; -
CC      EMBL; Z48542; CAA88428.1; -
CC      EMBL; AF002276; AAB93970.1; -
CC      PIR; S57911; S38508.
CC      HSP; P34034; 2LEU.
CC      InterPro; IPR002633; Bacteriocin II.
CC      Pfam; PF01721; Bacteriocin II; 1.
CC      ProDom; PD004452; Bacteriocin II; 1.
CC      TIGRFAMs; TIGR01847; bacteriocin_sig; 1.
KW      Antibiotic; Bacteriocin; Direct protein sequencing.
-FT
CHAIN   19 61      Bacteriocin sakacin P.
FT DISULFID 27 32      BY similarity.
SQ      SEQUENCE 61 AA; 6385 MW; B7BF14DCD28A73D CRC64;
* Query Match 21.1%; Score 64; DB 1; Length 61;
* Best Local Similarity 36.7%; Pred. No. 1.3;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

```

```

QY      3 YGTNGVHCTKNSLWGVKRLKMKYDQNTYMGRLQDILLGWATG 47
Db      20 YYG-NGVHCGRKHSCTDWMG-----TAIGNIGNNAANWATG 54

```

```

RESULT 5
Q8YK80
ID      Q8YK80      PRELIMINARY;      PRT;      590 AA.
AC      Q8YK80;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      All8037 protein.
GN      OrderedLocusNames=all8037;
OS      Anabaena sp. (strain PCC 7120).
OG      Plasmid pCC7120gamma.
OC      Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX      NCBI_TaxID=103690;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=2159285; PubMed=11759840;
RA      Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA      Watanabe A., Iiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA      Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA      Yasuda M., Tabata S.;
RT      "Complete genomic sequence of the filamentous nitrogen-fixing
RT      cyanobacterium Anabaena sp. strain PCC 7120.";
RL      DNA Rep. 8:205-213(2001).
DR      EMBL; AF003603; BAB77367.1; -
DR      PIR; AF2555; AF2555.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0009291; P:unidirectional conjugation; IEA.
DR      Pfam; PF02534; TraG; 1.
KW      Complete proteome; Plasmid.
SQ      SEQUENCE 590 AA; 64933 MW; 7ECA68DF6FE3B469 CRC64;

```

```

Query Match 20.7%; Score 63; DB 2; Length 590;
Best Local Similarity 35.3%; Pred. No. 21;
Matches 18; Conservative 8; Mismatches 23; Indels 2; Gaps 1;

```

```

QY      3 YGTNGVHCTKNSLWGVKRLKMKYDQNTYMGRLQDILLGWATGAFGKTF 53
Db      88 YIGT--PRGTKFVVGNGKRNITNPDKNRLYLPDVQRGVILVSGSGGKTF 136

```

```

RESULT 6
Q8RIM1
ID      Q8RIM1      PRELIMINARY;      PRT;      1582 AA.
AC      Q8RIM1;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Fusobacterium outer membrane protein family.
GN      OrderedLocusNames=FN1554;
OS      Fusobacterium nucleatum (subsp. nucleatum).
OC      Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC      Fusobacterium.
OX      NCBI_TaxID=76856;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=ATCC 25586;
RX      MEDLINE=21886394; PubMed=11889109;
RX      DOI=10.1128/JB.184.7.2005-2018.2002;
RA      Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA      Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA      Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA      Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA      Fonstein M., Kyrpides N.C., Overbeek R.;
RT      "Genome sequence and analysis of the oral bacterium Fusobacterium
RT      nucleatum strain ATCC 25586.";
RL      J. Bacteriol. 184:2005-2018(2002).
DR      EMBL; AB010464; AAL93680.1; -
DR      InterPro; IPR005546; Auto_transptbeta.
DR      InterPro; IPR001014; Ribosomal_L23.
DR      Pfam; PF03797; Autotransporter; 1.
DR      PROSITE; PS00050; RIBOSOMAL_L23; UNKNOWN_1.
KW      Complete proteome.
SQ      SEQUENCE 1582 AA; 167889 MW; 666C38783A00EB4F CRC64;

```

```

Query Match 20.7%; Score 63; DB 2; Length 1582;
Best Local Similarity 35.2%; Pred. No. 62;
Matches 19; Conservative 6; Mismatches 21; Indels 8; Gaps 3;

```

```

QY      4 YGTNGVHCTKNSLWGVKRLKMKYDQNTYMGRLQDILL---GWATGAFGKTF 53
Db      1310 FGTNGEYKTDTA--GVIDYKNHAY--GVAYVHENEIDKLGRGTGWTGIVHNTF 1359

```

```

RESULT 7
Q8RHH7
ID      Q8RHH7      PRELIMINARY;      PRT;      1630 AA.
AC      Q8RHH7;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Fusobacterium outer membrane protein family.
GN      OrderedLocusNames=FN2047;
OS      Fusobacterium nucleatum (subsp. nucleatum).
OC      Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC      Fusobacterium.
OX      NCBI_TaxID=76856;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=ATCC 25586;
RX      MEDLINE=21886394; PubMed=11889109;
RX      DOI=10.1128/JB.184.7.2005-2018.2002;
RA      Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA      Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA      Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA      Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA      Fonstein M., Kyrpides N.C., Overbeek R.;
RT      "Genome sequence and analysis of the oral bacterium Fusobacterium
RT      nucleatum strain ATCC 25586.";
RL      J. Bacteriol. 184:2005-2018(2002).
DR      EMBL; AB010507; AAL94131.1; -
DR      InterPro; IPR005546; Auto_transptbeta.
DR      Pfam; PF03797; Autotransporter; 1.

```

```
KW Complete proteome.
SQ SEQUENCE 1630 AA; 172741 MW; 38E9CE42F037B75A CRC64;

Query Match
Best Local Similarity 20.7%; Score 63; DB 2; Length 1630;
Matches 19; Conservative 6; Mismatches 21; Indels 8; Gaps 3;

QY 4 YGNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLODILL-----GWATGARGKTF 53
DB 1358 FGNGEYKDTDA--GVIDYKNHAY--GVAYVHNEIDIKRGIGWYTGIVHNTF 1407

RESULT 8
GLNA_METVO STANDARD; PRT; 446 AA.
AC P21154;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT- 05-JUL-2004 (Rel. 44, Last annotation update)
DE- Glutamine synthetase [EC 6.3.1.2] (Glutamate--ammonia ligase) (GS).
GN Name=glnA;
OS Methanococcus voltae.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=2188;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1537 / PS;
RX MEDLINE=90139872; PubMed=2575777; DOI=10.1016/0923-2508(89)90012-0;
RA Possot O., Sibold L., Aubert J.-P.;
RT "Nucleotide sequence and expression of the glutamine synthetase
structural gene, glnA, of the archaeobacterium Methanococcus voltae.";
RL Res. Microbiol. 140:335-371(1989).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
L-glutamine.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glutamine synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X53509; CAA37585.1; -.
DR PIR; A43995; A43995.
DR HSP; P06201; ILGR.
DR InterPro; IPR008147; Gln_synth_beta.
DR InterPro; IPR008146; Gln_synth_C.
DR InterPro; IPR004809; GlnA.
DR Pfam; PF001120; Gln-synth_N; 1.
DR Pfam; PF03951; Gln-synth_N; 1.
DR ProDom; PD001057; Gln_synth_C; 1.
DR TIGRFAMs; TIGR00653; GlnA; 1.
DR PROSITE; PS00180; GLNA 1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
KW Ligase.
SQ SEQUENCE 446 AA; 50199 MW; 95449E4DE8542690 CRC64;

Query Match
Best Local Similarity 20.6%; Score 62.5; DB 1; Length 446;
Matches 19; Conservative 9; Mismatches 17; Indels 11; Gaps 5;

QY 1 KTYVGTN--GVHCTKNSLW--GKVLKNMKYDQNTTYMGRLODILLGWATGARGKTF 52
DB 237 KPFFGMGSGWHNCQ--SIWLDGRPSF-----YDENNAH--QLSDICLSYICGILEHT 285

RESULT 9
Q8VM63 PRELIMINARY; PRT; 558 AA.
ID Q8VM63
```

```
AC Q8VM63;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cry39ORF2 protein.
GN Name=39orf2;
OC Bacillus thuringiensis (subsp. aizawai).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1433;
RN [1]
RP SEQUENCE FROM N.A.
RA Ito T., Sahara K., Bando H., Asano S.;
RT "Cloning and Expression of Novel Crystall Protein Genes cry39A and
39orf2 from Bacillus thuringiensis subsp. aizawai Buni-14 Encoding
Mosquitocidal Proteins.";
RL J. Insect Biotechnol. Sericulture 71:123-128(2002).
DR EMBL; AB074413; BAB72017.1; -.
SQ SEQUENCE 558 AA; 63360 MW; 26F6D5A013834207 CRC64;

Query Match
Best Local Similarity 20.6%; Score 62.5; DB 2; Length 558;
Matches 13; Conservative 8; Mismatches 17; Indels 9; Gaps 2;

QY 9 VHCTKNS-----LWGVKVR-LKNMKYDQNTTYMGRLODILLGWAT 46
DB 26 IECMSNEHSKSEEMMLWDEVKQAKQLSWRNLLYNGDFEDVSNWKT 72

RESULT 10
Q9PZ03 PRELIMINARY; PRT; 469 AA.
ID Q9PZ03
AC Q9PZ03;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF40.
GN Name=ORF40;
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum
granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=51677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99434230; PubMed=10502508; DOI=10.1006/viro.1999.9894;
RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
RL Virology 262:277-297(1999).
DR EMBL; AF162221; AAF05154.1; -.
DR HSP; P03956; 1CGL.
DR MEROPS; M10_032; -.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; P:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR PRINTS; PR006025; Pept_M_Zn_BS.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN 1.
DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN 1.
SQ SEQUENCE 469 AA; 53808 MW; 1BF4EB5C2101E241 CRC64;

Query Match
Best Local Similarity 20.2%; Score 61.5; DB 2; Length 469;
Matches 14; Conservative 9; Mismatches 7; Indels 9; Gaps 2;

QY 5 GTNGVHCTKNSLWGVRLKNMKYDQNTTYM--GRLODIL 41
DB 335 GDDEIATIRNLW-----YEYHKNSTLMNVRGVQDVL 366

RESULT 11
Q8AIJ6
```

<hr/>					
Q9DL93	ID	Q9DL93	PRELIMINARY;	PRT;	853 AA.
AC	Q9DL93;				
DT	01-MAR-2001	(TREMBLrel. 16, Created)			
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)			
DE	Envelope glycoprotein.				
GN	Name=env;				
OS	Human immunodeficiency virus 1.				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
NCBI_TaxID=11676;	[1]				
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=20541986; PubMed=11090186;				
RX	DOI=10.1128/JVI.74.24.11858-11872.2000;				
RA	Hu, Qx, Barry A.P., Wang, Zx, Connolly S.M., Peiper S.C.,				
RA	Greenberg M.L.;				
RT	"Evolution of the human immunodeficiency virus type 1 envelope during				
RT	infection reveals molecular correlates of specificity for coreceptor				
RT	utilization and AIDS pathogenesis.";				
J. Virol. 74:11858-11872(2000).	[2]				
RN	SEQUENCE FROM N.A.				
RP	Hu Q.-X., Barry A.P., Wang Z.-X., Connolly S.M., Goodman D.,				
RA	Peiper S.C., Greenberg M.L.;				
RA	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.				
RL	EMBL; AF310113; AAG36991.1; -.				
DR	HSP; P04578; 1DLB.				
DR	GO; GO:0016021; C:integral to membrane; IEA.				
DR	GO; GO:0019028; C:viral capsid; IEA.				
DR	GO; GO:0019031; C:viral envelope; IEA.				
DR	GO; GO:0005198; F:structural molecule activity; IEA.				
DR	InterPro; IPR000328; Env GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00317; GP41; 1.				
KW.	AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;				
KW	Transmembrane.				
SQ	SEQUENCE 853 AA; 96901 MW; 4BB4C83101B632E7 CRC64;				
<hr/>					
Query Match	20.2%;	Score 61.5;	DB 2;	Length 853;	
Best Local Similarity	35.9%;	Pred. No. 50;			
Matches 14;	Conservative 8;	Mismatches 14;	Indels 3;	Gaps	
<hr/>					
QY	7	NGVHTCKNSLGKVR---LKMKYDONTYMGRLQDILL 42			
Db	136	NNTSGTSSWGKVQEGEIKGCSFNITSIIRGVQEYSL 174			
<hr/>					
RESULT 14					
Q9DL79	ID	Q9DL79	PRELIMINARY;	PRT;	858 AA.
AC	Q9DL79;				
DT	01-MAR-2001	(TREMBLrel. 16, Created)			
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)			
DE	Envelope glycoprotein.				
GN	Name=env;				
OS	Human immunodeficiency virus 1.				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
NCBI_TaxID=11676;	[1]				
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=20541986; PubMed=11090186;				
RX	DOI=10.1128/JVI.74.24.11858-11872.2000;				
RA	Hu, Qx, Barry A.P., Wang, Zx, Connolly S.M., Peiper S.C.,				
RA	Greenberg M.L.;				
RT	"Evolution of the human immunodeficiency virus type 1 envelope during				
RT	infection reveals molecular correlates of specificity for coreceptor				
RT	utilization and AIDS pathogenesis.";				
J. Virol. 74:11858-11872(2000).	[2]				
RN	SEQUENCE FROM N.A.				
RP					

```
RA Hu Q.-X., Barry A.P., Wang Z.-X., Connolly S.M., Goodman D.,
RA Peiper S.C., Greenberg M.L.; to the EMBL/GenBank/DBJ databases.
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF310128; AAC37006.1; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR011010; DNA_brick_join_enz.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; C: cat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 858 AA; 97622 MW; B3D64720CF57DFAF CRC64;

Query Match 20.2%; Score 61.5; DB 2; Length 858;
Best Local Similarity 33.3%; Pred. No. 50;
Matches 13; Conservative 9; Mismatches 14; Indels 3; Gaps 1;

QY 7 NGVHCTKNSLWGKVR---LKNMKYDQNTTYMGRILQDILL 42
*  |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 136 NGHTGNTSSIWKEQKEIKNCSEFNITTGIRGKQVEYSL 174

RESULT 15
LCCA LEUGE . STANDARD; PRT; 61 AA.
AC P34036;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Bacteriocin leucocin A precursor (Leucocin A-UAL 187) (Leu A).
GN Name: lcnA;
OS Leuconostoc gelidium.
OG Plasmid pUG7.6.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1244;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-37.
RC STRAIN=UAL 187;
RX MEDLINE=92041660; PubMed=1840587;
RA Hastings J.W., Sailer M., Johnson K., Roy K.L., Vederas J.C.,
RA Stiles M.E.;
RT "Characterization of leucocin A-UAL 187 and cloning of the bacteriocin
RL gene from Leuconostoc gelidium.";
RN J. Bacteriol. 173:7491-7500(1991).
RP [2]
RP STRUCTURE BY NMR.
RC STRAIN=UAL 187;
RX MEDLINE=93120109; PubMed=8418850;
RA Sailer M., Helms G.L., Henkel T., Niemczura W.P., Stiles M.E.,
RA Vederas J.C.;
RT "15N- and 13C-labeled media from Anaerobacillus sp. for universal isotopic
RT labeling of bacteriocins: NMR resonance assignments of leucocin A from
RL Leuconostoc gelidium and nisin A from Lactococcus lactis.";
RL Biochemistry 32:310-318(1993).
RP [3]
RP STRUCTURE BY NMR.
RX MEDLINE=98060758; PubMed=9398233; DOI=10.1021/bi971263h;
RA Gallagher N.L.F., Sailer M., Niemczura W.P., Nakashima T.T.,
RA Stiles M.E., Vederas J.C.;
RT "Three-dimensional structure of leucocin A in trifluoroethanol and
RT dodecylphosphocholine micelles: spatial location of residues critical
RT for biological activity in type IIA bacteriocins from lactic acid
RL bacteria.";
RL Biochemistry 36:15062-15072(1997).
CC -!- FUNCTION: Inhibits a wide spectrum of lactic acid bacteria.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the bacteriocin class IIA/YNGV family.
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M64371; AAA68003.1; -.
DR PIR; A41657; A41657.
DR PDB; 1CW6; NMR; A=25-61.
DR PDB; 2LEU; NMR; @=25-61.
DR PDB; 3LEU; NMR; @=25-61.
DR InterPro; IPR002633; Bacteriocin_II.
DR InterPro; IPR010133; Bacteriocin_sig.
DR Pfam; PF01721; Bacteriocin_II; 1.
DR ProDom; PD004452; Bacteriocin_II; 1.
DR TIGRFAMs; TIGR01847; bacteriocin_sig; 1.
KW 3D-structure; Antibiotic; Bacteriocin; Direct protein sequencing;
KW Plasmid.
FT PROPEP 1 24 Bacteriocin leucocin A.
FT CHAIN 25 61
FT DISULFID 33 38
FT STRAND 26 26
FT TURN 29 30
FT STRAND 32 32
FT HELIX 42 53
FT TURN 54 54
SQ SEQUENCE 61 AA; 6586 MW; 4A645D631486409F CRC64;

Query Match 19.7%; Score 60; DB 1; Length 61;
Best Local Similarity 48.4%; Pred. No. 4.3;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

QY 3 YGTNGVHCTKNSL---WGKV-----RLKN 24
* |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 26 YYG-NGVHCTKSGCSVNWGEAFPSAGVHRLAN 55

Search completed: May 16, 2005, 17:34:00
Job time : 170 secs
```